



การพยากรณ์จำนวนนักท่องเที่ยวและอุณหภูมิในพื้นที่อุทยานแห่งชาติโดยแบบจำลอง

ผู้จัดทำ

เลขประจำตัว 64122250020 ชื่อ ณัฐติยา พรหมสันเทียะ

เลขประจำตัว 64122250025 ชื่อ วันวิสา โลหะชาติ

เสนอ

รองศาสตราจารย์ ดร.กัลยณัฐ กุหลาบเพชรทอง

รายงานนี้เป็นส่วนหนึ่งของวิชา CSD2301 วิทยาการข้อมูลเบื้องต้น

สาขาวิชาวิทยาการคอมพิวเตอร์และนวัตกรรมข้อมูล

คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสวนสุนันทา

ภาคเรียนที่ 1 ปีการศึกษา 2566

การพยากรณ์จำนวนนักท่องเที่ยวและอุณหภูมิในพื้นที่อุทยานแห่งชาติโดยแบบจำลอง

ณัฐติยา พรหมสันเทียะ และ วันวิสา โลหะชาติ

คณะวิทยาศาสตร์และเทคโนโลยี สาขาวิทยาการคอมพิวเตอร์และนวัตกรรมข้อมูล

บทคัดย่อ

ปัจจุบันนักท่องเที่ยวถือเป็นกลุ่มเป้าหมายสำคัญของพื้นที่อุทยานแห่งชาติของประเทศไทย ไม่ว่าจะเป็นนักท่องเที่ยวทั้งชาวไทยและชาวต่างชาติ ซึ่งนอกจากนักท่องเที่ยวแล้วยังมีสภาพอุณหภูมิของอากาศที่เป็นปัจจัยสำคัญในการเข้ามาท่องเที่ยวของนักท่องเที่ยวในพื้นที่อุทยานแห่งชาติอีกด้วย ดังนั้นการพยากรณ์การศึกษาจำนวนนักท่องเที่ยวและอุณหภูมิถือเป็นเรื่องที่น่าสนใจ เนื่องจากมีส่วนสำคัญต่อการวางแผนการท่องเที่ยวในพื้นที่อุทยานแห่งชาติในอนาคต งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบวิธีการพยากรณ์สำหรับการพยากรณ์จำนวนของนักท่องเที่ยวและอุณหภูมิในพื้นที่อุทยานแห่งชาติมี 2 วิธี คือ วิธีการถดถอยเชิงเส้น (Linear Regression) และ วิธีการคำนวณเพื่อนบ้านที่ใกล้ที่สุด (K-Nearest Neighbors) และก่อนที่จำใช้การเปรียบเทียบการพยากรณ์ทั้ง 2 วิธีนั้นได้มีการใช้วิธีการทำ Data Cleaning ในการเตรียมความพร้อมของข้อมูลเพื่อให้ข้อมูลมีคุณภาพสูงสุดก่อนที่จะนำมาใช้ในแบบจำลองโดยการจัดการข้อมูลที่สูญหาย (Missing Value) การตรวจสอบข้อมูลที่ผิดปกติ (Outliers) และ การตรวจสอบข้อมูลซ้ำ (Duplicate) หลังจากขั้นตอนการเตรียมชุดข้อมูลสำเร็จจึงได้นำข้อมูลมาสร้างแบบจำลองทั้ง 2 แบบเพื่อทำการพยากรณ์และวิเคราะห์ความแม่นยำของค่าเฉลี่ยของความคลาดเคลื่อนระหว่างค่าที่ทำนายได้จากโมเดลกับค่าที่จริงของข้อมูล (RMSE : Root Mean Squared Error) ค่าเฉลี่ยของความคลาดเคลื่อนที่มีค่าเป็นบวก ระหว่างค่าที่ทำนายได้จากโมเดลกับค่าจริงของข้อมูล (MAE : Mean Absolute Error) และ ความสามารถในการอธิบายของโมเดลที่ว่างจากความสับสน (variance) ในข้อมูล (R^2 : Coefficient of Determination) เมื่อวิเคราะห์และพิจารณาค่า RMSE MAE และ R^2 จากทั้งสองแบบจำลองสำหรับข้อมูลแล้วพบว่าแบบจำลองวิธีการถดถอยเชิงเส้นมีความแม่นยำและเป็นวิธีที่เหมาะสมที่สุดกับชุดข้อมูล

คำสำคัญ : การพยากรณ์, วิธีการถดถอยเชิงเส้น, วิธีการคำนวณเพื่อนบ้านที่ใกล้ที่สุด, Data Cleaning

Tourism and Temperature Prediction Model

Nuttiya Promsanthia and Wanwisa Lohachat

Faculty of Science and Technology, Computer Science and Data Innovation Branch

Abstract

Tourists are currently a key target group for Thailand's national parks, whether they are Thai or foreign visitors. Besides tourists, the temperature is also a crucial factor influencing tourist arrivals in national parks. Therefore, forecasting the number of tourists and temperature is an important study area for future tourism planning in national parks. This research aims to study and compare forecasting methods for predicting the number of tourists and temperature in national parks using two methods: Linear Regression and K-Nearest Neighbors (KNN). Prior to using these two methods, Data Cleaning was employed to prepare the data for modeling. This involved handling Missing Values, checking for Outliers, and removing Duplicate data. After successfully preparing the dataset, models were created using both methods to predict and analyze the accuracy of predictions. This was done by calculating the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2) to evaluate the models' performance. Upon analyzing and considering the RMSE, MAE, and R^2 values for both models, it was found that the Linear Regression method was the most accurate and suitable for the dataset.

Keywords: Forecasting, Linear Regression, K-Nearest Neighbors, Data Cleaning

บทนำ

การท่องเที่ยวในประเทศไทยเป็นอุตสาหกรรมที่มีบทบาทสำคัญต่อเศรษฐกิจ โดยการท่องเที่ยวไทยมีแนวโน้มเติบโตต่อเนื่อง ด้วยศักยภาพและโอกาสที่มีอยู่ในทรัพยากรธรรมชาติและวัฒนธรรมที่หลากหลายของประเทศไทย ซึ่งทำให้ประเทศไทยเป็นหนึ่งในจุดหมายปลายทางท่องเที่ยวยอดนิยมของนักท่องเที่ยวจากทั่วโลก

ซึ่งอุทยานแห่งชาติเป็นที่น่าสนใจสำหรับการท่องเที่ยวในประเทศไทย เนื่องจากมีความหลากหลายทางธรรมชาติที่สวยงาม ไม่ว่าจะเป็น อุทยานแห่งชาติเขาใหญ่ อุทยานแห่งชาติเขาสก อุทยานแห่งชาติเขาคิซมภูฏ หรือแม้แต่เขาใหญ่เขาหลวง ที่มีวิวทิวทัศน์ที่สวยงามได้เป็นอย่างดี นอกจากนี้ยังมีอุทยานทะเลที่สวยงาม เช่น อุทยานแห่งชาติหมู่เกาะสิมิลัน อุทยานแห่งชาติเกาะสิมิลัน และอุทยานแห่งชาติเกาะลันตา ที่มีทะเลใสทรายขาว และภูเขาที่ราบรินที่สร้างสถานที่ท่องเที่ยวที่น่าสนใจมากมายการเพิ่มจำนวนนักท่องเที่ยวให้มากขึ้นในแต่ละปี เป็นเป้าหมายที่รัฐบาลไทยมุ่งมั่นทำให้เกิดการสร้างสถานที่ท่องเที่ยวใหม่ๆ ที่น่าสนใจ ไม่ว่าจะเป็นสวนสนุก พิพิธภัณฑ์ และสถานที่ท่องเที่ยวที่ทำให้นักท่องเที่ยวต้องการมาเยี่ยมชม นอกจากนี้ยังมีการส่งเสริมการท่องเที่ยวในภูมิภาคต่างๆ ของประเทศ ไม่ว่าจะเป็นการสร้างที่พักผ่อน เช่น โรงแรม รีสอร์ท วิลล่า ที่มีความหรูหราและบรรยากาศที่น่าอยู่ หรือการส่งเสริมกิจกรรมท่องเที่ยวที่น่าสนใจในพื้นที่ เช่น เทศกาลวันสงกรานต์ งานประเพณีท้องถิ่น และกิจกรรมทางวัฒนธรรมอื่นๆ ที่ทำให้นักท่องเที่ยวได้สัมผัสวัฒนธรรมและทรัพยากรท้องถิ่นอย่างใกล้ชิดสร้างความหลากหลายในกิจกรรมท่องเที่ยวยังเป็นส่วนสำคัญในการดึงดูดนักท่องเที่ยว ไม่ว่าจะเป็นกิจกรรมเชิงวัฒนธรรม เช่น การเรียนรู้วัฒนธรรมท้องถิ่น การชมนิทรรศการศิลปะและวัฒนธรรม หรือกิจกรรมกลางแจ้ง เช่น ปีนเขา ทะเลสาบ และการผจญภัยต่างๆ ที่ทำให้นักท่องเที่ยวได้สัมผัสประสบการณ์ที่หลากหลายและน่าจดจำ

ดังนั้น การท่องเที่ยวในประเทศไทยไม่เพียงแต่เป็นการสนุกสนานและผ่อนคลาย แต่ยังเสริมสร้างรายได้ให้กับประเทศอีกด้วย ดังนั้นรัฐบาลไทยมีนโยบายในการส่งเสริมและพัฒนาอุตสาหกรรมท่องเที่ยวอย่างต่อเนื่อง เพื่อสร้างสังคมท่องเที่ยวที่ยั่งยืนและมีคุณภาพให้กับนักท่องเที่ยวทั้งในและต่างชาติที่มาเยี่ยมชมประเทศไทยในทุกๆ ช่วงเวลา

วัตถุประสงค์ของการวิจัย

1. สร้างแบบจำลองเพื่อพยากรณ์ค่าความแม่นยำในการทำนายจำนวนนักท่องเที่ยวและอุณหภูมิในพื้นที่อุทยานแห่งชาติ
2. เพื่อทราบความสัมพันธ์ระหว่างจำนวนนักท่องเที่ยวและอุณหภูมิในพื้นที่อุทยานแห่งชาติ

สมมติฐานของการวิจัย

1. การเปลี่ยนแปลงของจำนวนนักท่องเที่ยวที่มีต่อปริมาณอุณหภูมิตามสภาพแวดล้อมในพื้นที่อุทยานแห่งชาติ

2. การวิจัยแบบจำลองทางสถิติที่ใช้ในการวิเคราะห์ข้อมูลจำนวนนักท่องเที่ยวและอุณหภูมิในพื้นที่อุทยานแห่งชาติจะสามารถทำนายข้อมูลได้อย่างเหมาะสม

ขอบเขตของการวิจัย

1. ข้อมูลจำนวนนักท่องเที่ยวในแต่ละเดือนที่เข้าใช้บริการในพื้นที่อุทยานแห่งชาติของปี 2565

2. ข้อมูลปริมาณอุณหภูมิในแต่ละเดือนของพื้นที่อุทยานของปี 2565

วิธีการดำเนินการวิจัย

ขั้นตอนที่ 1 รวบรวมข้อมูล

รวบรวมข้อมูลรายชื่อสถานที่ท่องเที่ยวในพื้นที่อุทยานแห่งชาติของปี 2565 ข้อมูลจำนวนนักท่องเที่ยวในแต่ละเดือนที่เข้าใช้บริการในพื้นที่อุทยานแห่งชาติของปี 2565 และ ข้อมูลปริมาณอุณหภูมิในแต่ละเดือนของพื้นที่อุทยานของปี 2565

ขั้นตอนที่ 2 วิธีการดึงข้อมูล

นำเข้าข้อมูลที่ได้เก็บรวบรวมจาก กรมอุทยานแห่งชาติ สัตว์ป่า และ พันธุ์พืช สำนักอุทยานแห่งชาติ เป็นไฟล์นามสกุล .xlsx และทำการแปลงข้อมูลในตารางให้เป็นไฟล์นามสกุล .csv

Dataset

1. จำนวนนักท่องเที่ยวในแต่ละเดือนที่เข้าใช้บริการในพื้นที่อุทยานแห่งชาติของปี 2565

ที่มา : <https://catalog.dnp.go.th/dataset/stat-tourism>

ชื่อตัวแปร	คำอธิบาย
ID	ลำดับของสถานที่ท่องเที่ยวในพื้นที่อุทยานแห่งชาติ
อุทยานแห่งชาติ	ชื่อของสถานที่ท่องเที่ยวในพื้นที่อุทยานแห่งชาติ
สังกัด	สังกัดสำนักงานบริหารของแต่ละสถานที่
ปีงบประมาณ	ปีงบประมาณของข้อมูล
ต.ค.	จำนวนนักท่องเที่ยวเดือนตุลาคม 2565

พ.ย.	จำนวนนักท่องเที่ยวเดือนพฤศจิกายน 2565
ธ.ค.	จำนวนนักท่องเที่ยวเดือนธันวาคม 2565
ม.ค.	จำนวนนักท่องเที่ยวเดือนมกราคม 2565
ก.พ.	จำนวนนักท่องเที่ยวเดือนกุมภาพันธ์ 2565
มี.ค.	จำนวนนักท่องเที่ยวเดือนมีนาคม 2565
เม.ย.	จำนวนนักท่องเที่ยวเดือนเมษายน 2565
พ.ค.	จำนวนนักท่องเที่ยวเดือนพฤษภาคม 2565
มิ.ย.	จำนวนนักท่องเที่ยวเดือนมิถุนายน 2565
ก.ค.	จำนวนนักท่องเที่ยวเดือนกรกฎาคม 2565
ส.ค.	จำนวนนักท่องเที่ยวเดือนสิงหาคม 2565
ก.ย.	จำนวนนักท่องเที่ยวเดือนกันยายน 2565

2. ปริมาณอุณหภูมิในแต่ละเดือน ของพื้นที่อุทยานของปี 2565

ที่มา : <https://www.tmd.go.th/climate/summarymonthly>

ชื่อตัวแปร	คำอธิบาย
อุทยานแห่งชาติ	ชื่อของสถานที่ท่องเที่ยวในพื้นที่อุทยานแห่งชาติ
ต.ค.	ปริมาณอุณหภูมิในเดือนตุลาคม 2565
พ.ย.	ปริมาณอุณหภูมิในเดือนพฤศจิกายน 2565
ธ.ค.	ปริมาณอุณหภูมิในเดือนธันวาคม 2565
ม.ค.	ปริมาณอุณหภูมิในเดือนมกราคม 2565
ก.พ.	ปริมาณอุณหภูมิในเดือนกุมภาพันธ์ 2565
มี.ค.	ปริมาณอุณหภูมิในเดือนมีนาคม 2565
เม.ย.	ปริมาณอุณหภูมิในเดือนเมษายน 2565

พ.ค.	ปริมาณอุณหภูมิในเดือนพฤษภาคม 2565
มิ.ย.	ปริมาณอุณหภูมิในเดือนมิถุนายน 2565
ก.ค.	ปริมาณอุณหภูมิในเดือนกรกฎาคม 2565
ส.ค.	ปริมาณอุณหภูมิในเดือนสิงหาคม 2565
ก.ย.	ปริมาณอุณหภูมิในเดือนกันยายน 2565

Dataset ที่รวบรวมแล้ว ชุดข้อมูลที่รวบรวมแล้วนั้นได้มาจาก Dataset ที่หามาได้ 2 Dataset คือ จำนวนนักท่องเที่ยวในอุทยานแห่งชาติ ปีงบประมาณ พ.ศ. 2565 และ ข้อมูลอุณหภูมิแต่ละเดือนในปี 2565 โดยใช้ตัวแปรจากภาษาไทยเป็นตัวแปรภาษาอังกฤษและแสดงออกมาได้ชุดข้อมูล ที่บันทึกชื่อไฟล์ tourism_dataset.csv และ อธิบาย Features แต่ละ Features ของ Dataset ที่รวบรวมแล้วได้ดังนี้

ชื่อตัวแปร	คำอธิบาย
NationalPark	ชื่อของอุทยานแห่งชาติ 30 สถานที่
TempOct	อุณหภูมิโดยรวมในเดือนตุลาคมปี 2565 ของแต่ละพื้นที่อุทยาน 0 - 28.3 องศาเซลเซียส
TempNov	อุณหภูมิโดยรวมในเดือนพฤศจิกายนปี 2565 ของแต่ละพื้นที่อุทยาน 0 - 28.5 องศาเซลเซียส
TempDec	อุณหภูมิโดยรวมในเดือนธันวาคมปี 2565 ของแต่ละพื้นที่อุทยาน 0 - 26.4 องศาเซลเซียส
TempJan	อุณหภูมิโดยรวมในเดือนมกราคมปี 2565 ของแต่ละพื้นที่อุทยาน 0 - 27.6 องศาเซลเซียส
TempFeb	อุณหภูมิโดยรวมในเดือนกุมภาพันธ์ปี 2565 ของแต่ละพื้นที่อุทยาน 0 - 28.2 องศาเซลเซียส
TempMar	อุณหภูมิโดยรวมในเดือนมีนาคมปี 2565 ของแต่ละพื้นที่อุทยาน 0 - 30.0 องศาเซลเซียส
TempApr	อุณหภูมิโดยรวมในเดือนเมษายนปี 2565 ของแต่ละพื้นที่อุทยาน 0 - 29.8 องศาเซลเซียส
TempMay	อุณหภูมิโดยรวมในเดือนพฤษภาคมปี 2565 ของแต่ละพื้นที่อุทยาน 0 - 29.2 องศาเซลเซียส
TempJune	อุณหภูมิโดยรวมในเดือนมิถุนายนปี 2565 ของแต่ละพื้นที่อุทยาน 0 - 30.1 องศาเซลเซียส
TempJuly	อุณหภูมิโดยรวมในเดือนกรกฎาคมปี 2565 ของแต่ละพื้นที่อุทยาน 0 - 29.7 องศาเซลเซียส
TempAug	อุณหภูมิโดยรวมในเดือนสิงหาคมปี 2565 ของแต่ละพื้นที่อุทยาน 0 - 28.9 องศาเซลเซียส
TempSep	อุณหภูมิโดยรวมในเดือนกันยายนปี 2565 ของแต่ละพื้นที่อุทยาน 0 - 28.3 องศาเซลเซียส
TourismOct	จำนวนนักท่องเที่ยวในเดือนตุลาคมของแต่ละอุทยานแห่งชาติ 0 - 157662 คน

TourismNov	จำนวนนักท่องเที่ยวในเดือนพฤศจิกายนของแต่ละอุทยานแห่งชาติ 0 – 156093 คน
TourismDec	จำนวนนักท่องเที่ยวในเดือนธันวาคมของแต่ละอุทยานแห่งชาติ 0 - 250501 คน
TourismJan	จำนวนนักท่องเที่ยวในเดือนมกราคมของแต่ละอุทยานแห่งชาติ 0 - 158936 คน
TourismFeb	จำนวนนักท่องเที่ยวในเดือนกุมภาพันธ์ของแต่ละอุทยานแห่งชาติ 0 – 79772 คน
TourismMar	จำนวนนักท่องเที่ยวในเดือนมีนาคมของแต่ละอุทยานแห่งชาติ 0 - 66064 คน
TourismApr	จำนวนนักท่องเที่ยวในเดือนเมษายนของแต่ละอุทยานแห่งชาติ 0 - 109079 คน
TourismMay	จำนวนนักท่องเที่ยวในเดือนพฤษภาคมของแต่ละอุทยานแห่งชาติ 0 - 96790 คน
TourismJune	จำนวนนักท่องเที่ยวในเดือนมิถุนายนของแต่ละอุทยานแห่งชาติ 0 - 85674 คน
TourismJuly	จำนวนนักท่องเที่ยวในเดือนกรกฎาคมของแต่ละอุทยานแห่งชาติ 0 - 133876 คน
TourismAug	จำนวนนักท่องเที่ยวในเดือนสิงหาคมของแต่ละอุทยานแห่งชาติ 0 - 84645 คน
TourismSep	จำนวนนักท่องเที่ยวในเดือนกันยายนของแต่ละอุทยานแห่งชาติ 0 - 61213 คน
TourismTotal	ยอดรวมจำนวนนักท่องเที่ยวในปี 2565 ของแต่ละอุทยานแห่งชาติ 0 - 1428765 คน

ขั้นตอนที่ 3 จัดเตรียมข้อมูล

ในส่วนการเตรียมข้อมูลนั้นได้นำ Dataset ที่รวบรวมแล้วนั้นมาทำการ Data Cleaning เพื่อให้มีข้อมูลครบถ้วนในการนำไปคำนวณค่าของจำนวนนักท่องเที่ยวที่เข้ามาใช้บริการในอุทยานแห่งชาติเพื่อนำไปพยากรณ์ค่าปริมาณของนักท่องเที่ยวที่เข้ามาใช้บริการพื้นที่อุทยานแห่งชาติและอุณหภูมิแต่ละเดือนในปี 2565 ซึ่ง Dataset ก่อนที่จะมีการทำ Data Cleaning มีข้อมูลภายในชุดข้อมูล

1. Data Cleaning มี 3 วิธีดังนี้

1.1 Missing Value

1.2 Outlier

1.3 Duplicate

2. Transform Data โดยแปลงข้อมูลจากข้อมูลประเภทตัวอักษรหรือตัวเลขที่อยู่ปะปนกันให้เป็นตัวเลขเพื่อให้คอมพิวเตอร์นำไปประมวลผลได้

ขั้นตอนที่ 4 แบ่งข้อมูล

แบ่งข้อมูลออกเป็นชุดสำหรับใช้ในการฝึกและทดสอบชุดข้อมูลด้วยแบบจำลอง โดยการสุ่มเลือกข้อมูลที่เหมาะสมสำหรับหารฝึกและทดสอบด้วยสัดส่วนในการแบ่งข้อมูลคือ 80-20

ขั้นตอนที่ 5 สร้างและฝึกทดสอบแบบจำลอง

1. สร้างแบบจำลอง Linear regression โดยใช้ชุดข้อมูลการฝึกทดสอบ
2. สร้างแบบจำลอง K-Nearest Neighbors (KNN) โดยใช้ชุดข้อมูลการฝึกทดสอบ
3. ทำการฝึกทดสอบแบบจำลอง Linear regression และ K-Nearest Neighbors (KNN) ด้วยชุดข้อมูลฝึกทดสอบ

ขั้นตอนที่ 6 ทดสอบแบบจำลองและเปรียบเทียบ

ใช้ชุดข้อมูลสำหรับการฝึกทดสอบเพื่อประเมินประสิทธิภาพความแม่นยำในการทำนายของทั้งสองแบบจำลองโดยคำนวณค่าของ RMSE MAE และ R^2 เพื่อเปรียบเทียบความแม่นยำของแบบจำลองที่ใช้

ขั้นตอนที่ 7 ผลลัพธ์และสรุปผลการวิจัย

สรุปผลลัพธ์ของการวิจัยในการพยากรณ์จำนวนของนักท่องเที่ยวและอุณหภูมิในแต่ละเดือนของพื้นที่อุทยาน โดยความสามารถของ แบบจำลอง Linear regression และ K-Nearest Neighbors (KNN) ในการพยากรณ์ และ ระบุความเหมาะสมการใช้แบบจำลองที่กำลังวิเคราะห์

ผลการวิจัย

Model	RMSE	MAE	R^2
Linear regression	8227.78	5836.37	-1.0142
K-Nearest Neighbors (KNN)	9508.96	6293.37	-1.9233

จากตาราง สรุปได้ว่า

1. ความสามารถในการพยากรณ์

แบบจำลอง Linear regression เหมาะสมสำหรับการนำมาพยากรณ์ค่าปริมาณของนักท่องเที่ยวที่เข้ามาใช้บริการพื้นที่อุทยานแห่งชาติและอุณหภูมิในแต่ละเดือนในปี 2565 โดย ค่า RMSE มีค่าเท่ากับ 8227.77858881263 มีค่าสูงแต่ค่าที่น้อยกว่าค่า RMSE ของ K-Nearest Neighbors (KNN) ที่มีค่าเท่ากับ 9508.95963497822 ซึ่งมีค่าเป็นจำนวนเต็มบวกแสดงถึงความแตกต่างเฉลี่ยระหว่างค่าที่ทำนายได้กับค่าจริง

ของจำนวนนักท่องเที่ยวแสดงว่าโมเดลมีความคลาดเคลื่อนมากขึ้น, ค่า MAE มีค่าเท่ากับ 5836.36593081287, ค่า R^2 (อัตราความแม่นยำ) มีค่าเท่ากับ -1.014171290407619 ซึ่งมีค่าน้อยกว่าค่า MAE และ ค่า R^2 (อัตราความแม่นยำ) ใน K-Nearest Neighbors (KNN)

2. ความเหมาะสมในการใช้แบบจำลอง

แบบจำลอง Linear regression เป็นโมเดลที่ทำได้ดีกว่าการใช้ แบบจำลอง K-Nearest Neighbors (KNN) ในการพยากรณ์จำนวนของนักท่องเที่ยวที่เข้ามาใช้บริการพื้นที่อุทยานแห่งชาติตามอุณหภูมิในแต่ละเดือนของปี 2565

ดังนั้น การใช้แบบจำลอง Linear regression เป็นการเลือกที่เหมาะสมในการพยากรณ์จำนวนนักท่องเที่ยวและอุณหภูมิในพื้นที่อุทยานแห่งชาติโดยแบบจำลอง

สรุปผลการวิจัย

1. ความสัมพันธ์ระหว่างจำนวนนักท่องเที่ยว และ อุณหภูมิแต่ละเดือน

ผลการวิจัยจึงสรุปได้ว่ามีความสัมพันธ์ระหว่างจำนวนนักท่องเที่ยว และ อุณหภูมิแต่ละเดือนอย่างชัดเจน โดยผลของการวิเคราะห์ของข้อมูลในปี 2565 เป็นตัวชี้วัดในการระบุแนวโน้มความสัมพันธ์ของทั้งสองตัวแปรนี้

2. การสร้างแบบจำลองการพยากรณ์ทางสถิติที่ใช้ในการวิเคราะห์ข้อมูลจำนวนนักท่องเที่ยวและอุณหภูมิของแต่ละเดือนได้แก่ แบบจำลอง Linear regression และ แบบจำลอง K-Nearest Neighbors (KNN) ผลการทดสอบแบบจำลองทั้งสองแสดงให้เห็นว่า Linear regression มีความแม่นยำกว่าในการทำนายข้อมูลโดยมีค่า Root Mean Squared Error (RMSE) / Mean Absolute Error (MAE) และ R-squared: Coefficient of Determination (R^2) ที่ต่ำกว่าในแบบจำลอง K-Nearest Neighbors (KNN) ซึ่งแสดงให้เห็นว่า แบบจำลอง Linear regression มีความเหมาะสมที่สูงกว่าในการใช้ทำนายจำนวนนักท่องเที่ยวจากอุณหภูมิในแต่ละเดือนของปี 2565

Program

<https://github.com/Nnuttiya/PredictionTourism2022.git>

เอกสารอ้างอิง

การแบ่งกลุ่มข้อมูล แบบค่ามาตรฐานเคมีน (K-means algorithm)

<https://wjst.wu.ac.th/index.php/wuresearch/article/view/7093>

การพยากรณ์จำนวนนักท่องเที่ยวชาวจีนที่เดินทางเข้ามาในประเทศไทย

<https://opac01.rbru.ac.th/multim/journal/03502.pdf>

การพยากรณ์จำนวนนักท่องเที่ยวต่างชาติที่เดินทางมาท่องเที่ยวในประเทศไทยโดยวิธีบอกซ์-เจนกินส์และวิธีการของวินเตอร์

<https://li01.tci-thaijo.org/index.php/tstj/article/view/15129>

ปัจจัยที่มีอิทธิพลต่อการท่องเที่ยวอุทยานแห่งชาติภูกระดึงจังหวัดเลยของนักท่องเที่ยวชาวไทย

http://www.ba-abstract.ru.ac.th/AbstractPdf/2562-5-13_1599637611.pdf

แบบจำลองทางคณิตศาสตร์ พยากรณ์ จำนวนนักท่องเที่ยวในอุทยาน ประวัติศาสตร์พนมรุ้ง

<https://ph02.tci-thaijo.org/index.php/scibru/article/view/242108>

พฤติกรรมและแรงจูงใจของนักท่องเที่ยวชาวไทยที่เดินทางมาท่องเที่ยวอุทยานแห่งชาติเขาใหญ่ อำเภอปากช่อง จังหวัดนครราชสีมา

http://dspace.bu.ac.th/bitstream/123456789/4201/3/rungfa_saka.pdf

ความพึงพอใจของนักท่องเที่ยวที่มีต่อสิ่งอำนวยความสะดวกและการให้บริการ ของอุทยานแห่งชาติภูผาเทิบ จังหวัดมุกดาหาร

<https://wjst.wu.ac.th/index.php/wuresearch/article/view/6925>

สภาพภูมิอากาศกับการเปลี่ยนแปลงของสภาพแวดล้อม

<http://dspace.spu.ac.th/handle/123456789/3473?mode=full>

การเปลี่ยนแปลงสภาพภูมิอากาศกับอุตสาหกรรมการท่องเที่ยวไทย

<https://www.pier.or.th/abridged/2023/24/>